Overlap and activity glass transitions in plaquette spin models with hierarchical dynamics

Robert M. Turner,[1] Robert L. Jack,[2] and Juan P. Garrahan[1]
[1]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*
[2]*Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom*
(Received 21 April 2015; revised manuscript received 6 July 2015; published 10 August 2015)

We consider thermodynamic and dynamic phase transitions in plaquette spin models of glasses. The thermodynamic transitions involve coupled (annealed) replicas of the model. We map these coupled-replica systems to a single replica in a magnetic field, which allows us to analyze the resulting phase transitions in detail. For the triangular plaquette model (TPM), we find for the coupled-replica system a phase transition between high- and low-overlap phases, occurring at a coupling $\varepsilon^*(T)$, which vanishes in the low-temperature limit. Using computational path sampling techniques, we show that a single TPM also displays "space-time" transitions between active and inactive dynamical phases. These first-order dynamical transitions occur at a critical counting field $s_c(T) \gtrsim 0$ that appears to vanish at zero temperature in a manner reminiscent of the thermodynamic overlap transition. In order to extend the ideas to three dimensions, we introduce the square pyramid model, which also displays both overlap and activity transitions. We discuss a possible common origin of these various phase transitions, based on long-lived (metastable) glassy states.

DOI: 10.1103/PhysRevE.92.022115 PACS number(s): 64.60.−i, 64.70.Q−, 75.10.Hk, 05.40.−a

I. INTRODUCTION

As supercooled liquids approach their glass transitions, one observes a very sharp increase in their viscosities and structural relaxation times. The physical mechanism underlying this slow dynamics remains controversial [1–4]. Some theories, particularly the random first-order transition (RFOT) theory [5], propose that glassy systems are approaching some kind of thermodynamic phase transition, with associated collective (slow) dynamics. The existence of such phase transitions can be probed by computing the free energy of a pair of coupled copies (or replicas) of the system and searching for a transition as a function of both temperature and coupling strength. These transitions link an equilibriumlike phase where the replicas are different from each other (the liquid) to one where they become very similar (the glass) [6]. The similarity between the configurations is measured by an overlap variable, which is the order parameter for the transition. An alternative approach, that of dynamical facilitation [7], links glassy behavior to a dynamical "space-time" phase transition. This transition is explored through distributions of time-integrated observables, which quantify activity in the dynamics [8,9]. Based on these distributions, one may infer the existence of transitions between an active dynamical phase (the equilibrium liquid) and an inactive phase (the nonequilibrium glass). The order parameter for these transitions is the dynamical activity.

In this work, we investigate plaquette spin models of glasses [10], for which both overlap fluctuations and dynamical activity fluctuations can be analyzed, by a combination of analytical and computational methods. We concentrate on two models, whose relaxation behavior is similar to that of the facilitated East model [11–13]; their relaxation times increase faster than an Arrhenius law at low temperatures, but the equilibrium relaxation time is finite at all positive temperatures, diverging only as $T \to 0$. We present evidence that these models support *both* dynamic and thermodynamic phase transitions. In the thermodynamic case we consider a coupling between two *annealed* replicas, and transitions *occur only for nonzero (positive) values of the coupling*.

We argue that these results provide a connection between the (apparently quite different) "thermodynamic" and "dynamic" theories of the glass transition. This connection is built on the idea of metastability, which is intrinsically connected to glassy behavior. The formation of a metastable state in a finite-dimensional system requires that small perturbations in that state do not grow: The system prefers to relax back into the metastable state. This stability to small perturbations may be described in terms of an interfacial cost that acts to penalize local perturbations. Different theories ascribe different origins to these interfacial costs, which might be either static or dynamic, depending on the system of interest and the kinds of fluctuation being considered. However, the existence of these interfacial costs seems quite generic and may be useful for rationalizing different kinds of phase transition in these systems.

The main results of this work are as follows. We analyze the triangular plaquette model (TPM) in two spatial dimensions [14,15] and a three-dimensional variant of this model, which we refer to as the square pyramid model (SPyM). In Sec. II, we show that two (annealed) coupled replicas of these systems can be mapped to a single replica in a magnetic field, and we derive some useful features of this single-replica system, which place constraints on the kind of phase transitions that can occur. Some of these results were derived in previous work [16–18], but our analysis contains several new insights. In Sec. III, we show numerical evidence that the TPM in a magnetic field supports a phase transition in the two-dimensional (2D) Ising universality class. It then follows from the mappings in Sec. II that the coupled replicas of these systems also support a similar phase transition. In Sec. IV, we show that the TPM also supports dynamical space-time phase transitions, similar to those in [8,9,19]. In Sec. V, we introduce the SPyM and show evidence that it supports phase transitions in the coupled-replica setting and dynamical space-time phase transitions. Finally, in Sec. VI, we discuss the relationships between the thermodynamic and the dynamical phase transitions that we have found, and we consider the consequences of these results for theories of the glass transition.

II. PLAQUETTE MODELS, COUPLED REPLICAS, AND MAPPING TO A SYSTEM IN A FIELD

A. Models

We consider plaquette spin models defined in terms of classical Ising spins on regular lattices, with energy functions of the form

$$E_J(\sigma) \equiv -\frac{J}{2}\sum_{\mu}\sigma_{i_\mu}\sigma_{j_\mu}\cdots\sigma_{k_\mu}, \tag{1}$$

where $\sigma_i = \pm 1$, with i indicating a lattice site ($i = 1,\ldots,N$), and where the interactions are in terms of products of spins $\sigma_{i_\mu}\sigma_{j_\mu}\cdots\sigma_{k_\mu}$ around the *plaquettes* μ of the lattice. See [10,14,15] for a more general overview of the relevant properties of these systems. On a square lattice, one labels each square plaquette with an index μ, and $\{\sigma_{i_\mu},\sigma_{j_\mu},\ldots,\sigma_{k_\mu}\}$ is the set of four spins on the vertices of plaquette μ. This construction is easily generalized to higher dimensions: For a cubic lattice and cubic "plaquettes," each term in the energy would involve eight spins. This motivates us to define plaquette variables $\tau_\mu = \sigma_{i_\mu}\sigma_{j_\mu}\cdots\sigma_{k_\mu}$.

An interesting model in this class is the TPM [10], where the lattice is triangular and the interactions are between triplets of spins in the corners of upward-pointing triangles,

$$E_J(\sigma) \equiv -\frac{J}{2}\sum_{\mu=\triangle}\sigma_{i_\mu}\sigma_{j_\mu}\sigma_{k_\mu} \quad \text{(TPM)}. \tag{2}$$

The geometrical setting is shown in Fig. 1(a). Our analysis rests on a correspondence between configurations of the spin variables σ_i and the plaquette variables τ_μ. If we first consider rhombus-shaped systems whose linear size is an integer power of 2, with periodic boundaries, then there is a one-to-one mapping between spin configurations and plaquette configurations. (It is clear that every spin configuration corresponds to a single plaquette configuration, but the existence of a spin configuration corresponding to *every* plaquette configuration is less trivial [10,12,14,15].) For systems of different sizes or with different boundary conditions, the correspondence is not perfectly one to one, but these deviations turn out to be irrelevant in the thermodynamic limit. In Sec. V below, we also discuss the SPyM, a 3D model with the same one-to-one correspondence, on the body-centered cubic (bcc) lattice.

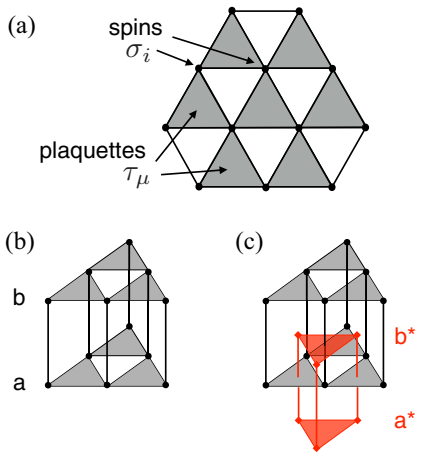

FIG. 1. (Color online) (a) Illustration of spins and plaquettes in the TPM. The spins σ_i are located on the vertices of the lattice. The plaquette variables τ_μ are located on the upward-pointing triangles (shaded). Each plaquette is associated with three spins and the variable τ_μ is given by the product of these spins. (b), (c) Geometrical illustration of the duality relation (14) in the TPM. Panel (b) shows two TPM systems, a and b, with coupling as in (3). Panel (c) shows the location of the sites of the dual problem, again two coupled TPMs a^* and b^*. The plaquettes in the dual system bisect the coupling interactions in the direct system, and vice versa.

In cases where the one-to-one mapping holds exactly, the fully polarized state $\sigma_i = 1\ \forall\ i$ is the unique ground state of (1). In terms of the plaquette variables, the ground state is $\tau_\mu = 1\ \forall\ \mu$, and the elementary excitation is a "defect," $\tau_\mu = -1$. Two-body spin correlations vanish in these models [10], although higher-order spin correlations are finite and allow access to a growing length scale at low temperatures [20]. Also, since there is a one-to-one mapping between spins and plaquettes, the thermodynamic properties of these models are those of noninteracting binary plaquette variables [10,12,14,15] or a free gas of "defective" plaquettes (with $\tau_\mu = -1$) [21].

However, while the thermodynamic properties of plaquette models are trivial, their (single spin-flip) dynamics is not. This effect arises because flipping a single spin σ_i changes the states of all of the plaquettes in which it participates. The plaquette dynamics is therefore "kinetically constrained" [10,12,15,22], possibly leading to complex glassy dynamics at low temperatures. This is what occurs, for example, in the TPM whose dynamical properties are similar to those of the East facilitated model [10,12,15], displaying "parabolic" super-Arrhenius relaxation, dynamic heterogeneity, and other characteristic features of the glass transition [4].

B. Coupled replicas

To probe thermodynamic overlap fluctuations, we consider two coupled replicas of a plaquette model [6,18,23–25]. The energy function of the combined system is

$$E_{J,\varepsilon}(\sigma^a,\sigma^b) \equiv E_J(\sigma^a) + E_J(\sigma^b) - \varepsilon\sum_i \sigma_i^a\sigma_i^b, \tag{3}$$

where σ^a and σ^b are the spin configurations in the replicas a and b. The overlap,

$$Q(\sigma^a,\sigma^b) \equiv \sum_i \sigma_i^a\sigma_i^b, \tag{4}$$

measures how similar the two copies are, and the strength of their coupling is given by its conjugate field ε. The coupling (3) is denoted *annealed* since both replicas are allowed to fluctuate on an equal footing. The case of *quenched* coupling, in contrast, involves one of the replicas being frozen in an equilibrium configuration. Here we only consider the case of annealed coupling, which is easier to treat both analytically and numerically. Hence, the partition function for these two coupled replicas is

$$Z_2(J,\varepsilon) = \sum_{\sigma^a,\sigma^b} e^{-\beta E_{J,\varepsilon}(\sigma^a,\sigma^b)}, \tag{5}$$

where the sum runs over the configurations σ^a,σ^b, that is, over all $\sigma_i^a = \pm 1$ and all $\sigma_i^b = \pm 1$. Here and in the following, we

sometimes set $\beta = 1$ where there is no ambiguity [for example, the left hand side of (5) should strictly be $Z_2(\beta J, \beta\varepsilon)$, but we suppress the dependence on β for simplicity].

C. Mapping to a single system in a field

In [18], a mapping was derived between the free energies of the coupled system (5) and a single plaquette model in a magnetic field. Here we present a mapping between (sets of) configurations of these systems, which extends that analysis, as well as recovering the same mapping between free energies.

We introduce overlap variables $q_i = \sigma_i^a \sigma_i^b$ on each site: Our aim is to calculate the statistical weight of a particular configuration of these variables. This weight is

$$W_2(q|J,\varepsilon) = \sum_{\sigma^a,\sigma^b} e^{-\beta E_{J,\varepsilon}(\sigma^a,\sigma^b)} \prod_i \delta\left(q_i - \sigma_i^a \sigma_i^b\right). \quad (6)$$

We now perform the sum over the σ variables. If we sum over σ^b first, we obtain

$$W_2(q|J,\varepsilon) = \sum_{\sigma^a} \exp\left[\frac{\beta J}{2} \sum_\mu \sigma_{i_\mu}^a \sigma_{j_\mu}^a \cdots \sigma_{k_\mu}^a \times (1 + q_{i_\mu} q_{j_\mu} \cdots q_{k_\mu}) + \beta\varepsilon \sum_i q_i\right].$$

For the summation over σ^a we replace $\sigma_{i_\mu}^a \sigma_{j_\mu}^a \cdots \sigma_{k_\mu}^a$ with τ_μ^a. Then we use the characteristic feature of the model, that plaquette and spin configurations are in a one-to-one correspondence, so we replace the sum over the σ_i^a with a sum over the τ_μ^a,

$$W_2(q|J,\varepsilon) = \sum_{\tau^a} \exp\left[\frac{\beta J}{2} \sum_\mu \tau_\mu^a (1 + q_{i_\mu} q_{j_\mu} \cdots q_{k_\mu}) + \beta\varepsilon \sum_i q_i\right].$$

Performing the sum, we arrive at

$$W_2(q|J,\varepsilon) = (4\cosh\beta J)^{N/2} e^{-\beta E_{J'}(q) + \beta\varepsilon \sum_i q_i}, \quad (7)$$

with

$$\beta J' = \ln\cosh(\beta J). \quad (8)$$

We recognize the exponential term in (7) as the statistical weight of a configuration $\sigma = q$ for a single plaquette model with energy scale J' in a magnetic field $h = \epsilon$.

To explore the consequences of this property for the free energy, we observe $Z_2(J,\varepsilon) = \sum_q W_2(q|J,\varepsilon)$, so that

$$Z_2(J,\varepsilon) = (4\cosh\beta J)^{N/2} Z_1(J',\varepsilon), \quad (9)$$

where

$$Z_1(J,h) = \sum_\sigma e^{-\beta E_J(\sigma) + \beta h \sum_i \sigma_i} \quad (10)$$

is the partition function of a single plaquette model in a field h. In addition, this latter system is known to have an exact duality [16,17],

$$Z_1(J,h) = (\sinh\beta J \sinh 2\beta h)^{N/2} Z_1(\tilde{J},\tilde{h}), \quad (11)$$

where

$$e^{-\beta\tilde{J}} = \tanh(\beta h), \quad e^{-2\beta\tilde{h}} = \tanh(\beta J/2). \quad (12)$$

From Eqs. (9)–(12) the duality of the coupled plaquette system follows

$$Z_2(J,\varepsilon) = (\sinh\beta J \sinh\beta\varepsilon)^N Z_2(J^*,\varepsilon^*), \quad (13)$$

with

$$e^{-\beta\varepsilon^*} = \tanh(\beta J/2), \quad e^{-\beta J^*} = \tanh(\beta\varepsilon/2). \quad (14)$$

This duality is precisely the one obtained in [18] for the two coupled replicas of the TPM. (Note that if $\tanh y = e^{-2x}$, then $\tanh x = e^{-2y}$, which follows from the definition of the tanh function and facilitates inversion of these duality transforms.)

D. Dualities and phase transitions

The mapping from two coupled replicas to a single system in a field has useful consequences, since we may exploit existing results for plaquette models in magnetic fields. The duality of this model, Eq. (11) (see also [17]), implies a duality relation for the free energy $F_1 = -T \ln Z_1$,

$$\beta F_1(h,J) + \frac{N}{2}\ln\sinh(2\beta h) = \beta F_1(\tilde{h},\tilde{J}) + \frac{N}{2}\ln\sinh(2\beta\tilde{h}), \quad (15)$$

where $\tilde{h}$ and $\tilde{J}$ are given in Eq. (12). Phase transitions appear as singularities in the free energy density $f_1 = \lim_{N\to\infty} F_1/N$.

From (15), if F_1 is singular at (h,J), it is also singular at $(\tilde{h},\tilde{J})$. This places constraints on the possible phase behavior of the system. In particular, if the system supports only a single phase transition, it must occur for parameters such that $(h,J) = (\tilde{h},\tilde{J})$. This condition defines a line in the (h,J) plane, which is given by

$$\beta J = -\ln\tanh\beta h. \quad (16)$$

On this line one has also $\sinh\beta J \sinh 2\beta h = 1$. Phase transitions that occur on such lines were investigated in [16,17]: The plaquette models considered there support a single critical point that occurs at some point (J_c,h_c) on this line, with first-order phase coexistence occurring on the part of the line with $J > J_c$.

We note that (15) resembles the Kramers-Wannier duality of the Ising model, which allows the position of the critical point to be identified exactly in that model. Here the situation is different because the transition takes place at finite h, in contrast to the Ising transition, which is known to take place at zero field, by symmetry. For this reason, the duality of the TPM does not fully determine the position of the critical point, but restricts it to the line (16) within the (h,J) plane.

From the above mapping, the phase transitions of the TPM in a field correspond to phase transitions in the coupled-replica system: The first-order transition line separates a state with low overlap (small ε) from one with high overlap (large ε). For the coupled replicas, the self-dual line is

$$\sinh(\beta J)\sinh(\beta\varepsilon) = 1. \quad (17)$$

This situation, where the self-dual line for the coupled-replica system contains a first-order transition region and a critical

point, was proposed for the TPM in Ref. [18]. We present numerical evidence for this situation in Sec. III below.

E. Other consequences of dualities and symmetries

In this section, we explore some further consequences of the results derived thus far. First, we note that the relation (7) means that for a coupled-replica system at parameters (J,ε), the probability of a particular configuration of the overlap variables q is the same as the probability of finding the configuration $\sigma = q$ for a single system in a field, with parameters $(J',h = \varepsilon)$. From a numerical perspective, the single system in a field is much simpler to simulate, and the result (7) means that such a simulation provides direct access to all observables based on the overlap variables. (This result is much stronger than a mapping at the level of free energies.)

Second, for a geometrical interpretation of the duality relation (14), we refer to Figs. 1(b) and 1(c). The original coupled system can be thought of as a lattice consisting of two parallel layers, a and b. The duality relation (14) may be interpreted as a mapping between two different two-layer systems, where the plaquette energy scale in one model determines the interlayer coupling in the other, and vice versa. Figure 1(c) illustrates this situation, in which the interlayer "bonds" in the original system intersect the intralayer plaquettes in the dual system, and vice versa. This geometrical way of seeing the duality easily generalizes to other lattices and plaquette interactions.

Third, the duality relation for a plaquette model in a field can be used to analyze the behavior of its free energy in the vicinity of a (presumed) critical point. We assume that a critical point exists somewhere on the self-dual line and that this critical point is in the Ising universality class, as is found generically [16] (see also below). The free energy is singular at the critical point (h_c,J_c), and the form of this singularity is universal. Given (15), it is convenient to define

$$\beta F_{\rm symm}(h,J) = \beta F_1(h,J) + \frac{N}{2}\ln\sinh(2\beta h). \tag{18}$$

The singular behavior of $F_{\rm symm}$ is the same as that of F_1 since the added term is regular. The reason for introducing $F_{\rm symm}$ is that the duality relation (15) now reads simply

$$F_{\rm symm}(h,J) = F_{\rm symm}(\tilde{h},\tilde{J}). \tag{19}$$

To investigate the universal behavior of $F_{\rm symm}$, we introduce fields $\mathcal{H} = \mathcal{H}(h,J)$ and $\mathcal{J} = \mathcal{J}(h,J)$. These fields are defined to be equal to zero at the critical point and will correspond to the directions in parameter space that are relevant under the renormalization group. The universal character of phase transitions means that these fields can be chosen such that

$$F_{\rm symm}(h,J) - F_{\rm symm}(h_c,J_c) \simeq a_0[F_I(\mathcal{H},\mathcal{J}) - F_I(0,0)], \tag{20}$$

where a_0 is a constant and $F_I(\mathcal{H},\mathcal{J})$ is the free energy of an Ising model in a field $h_I = \mathcal{H}$ and with coupling $J_I = J_{c,I} + \mathcal{J}$, in which $J_{c,I}$ is the critical coupling of the Ising model. The approximate equality in (20) accounts for nonuniversal contributions which are negligible for sufficiently large systems and close enough to the critical point.

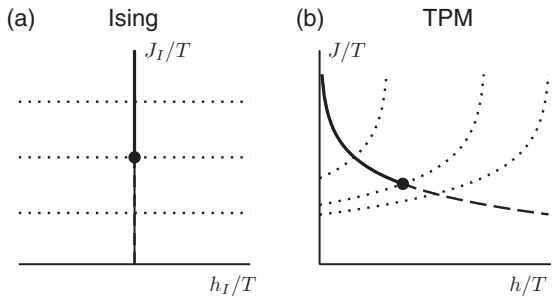

FIG. 2. Illustration of the relation between critical behavior of the Ising model and the TPM. (a) Ising phase diagram. The $h_I = 0$ axis is a symmetry line; there is a critical point indicated by a circle, with a first-order (phase coexistence) line for large J_I, indicated by a solid line. Selected lines of constant J_I are indicated by dotted lines. (b) The corresponding situation for the TPM in a field. On the solid-and-dashed line (16), the system has a discrete (Z_2) symmetry: A critical point and phase coexistence both occur on this line, as indicated. The dotted lines are obtained from (22) for three different values of h_0 and correspond to the lines of constant J_I in panel (a). Near the critical point, they indicate the direction of the most relevant renormalization group flow.

The symmetry of the Ising model under inversion of the field h_I means that $F_I(\mathcal{H},\mathcal{J}) = F_I(-\mathcal{H},\mathcal{J})$. The corresponding symmetry relation for the TPM is (19), and the operation of this symmetry should invert the magnetic field $\mathcal{H}$ but preserve the coupling $\mathcal{J}$ in the equivalent Ising model. Hence, the functions $\mathcal{H}$ and $\mathcal{J}$ should satisfy

$$\mathcal{H}(\tilde{h},\tilde{J}) = -\mathcal{H}(h,J), \qquad \mathcal{J}(\tilde{h},\tilde{J}) = \mathcal{J}(h,J). \tag{21}$$

A geometrical interpretation of (20) is shown in Fig. 2. The self-dual line (16) corresponds to the symmetry line $\mathcal{H} = 0$, which corresponds to $h_I = 0$ for the Ising model. We now derive TPM analogs of the dotted lines $J_I = \text{const.}$ in Fig. 2(a). That is, we seek curves $J(h)$ such that $\mathcal{J}(h,J(h))$ is constant.

In fact, since (20) applies only close to the critical point, it does not fully fix the dependence of $\mathcal{H}$ and $\mathcal{J}$ on (h,J). However, a consistent choice for the family of curves with constant $\mathcal{J}$ is

$$\beta J_{h_0}(h) = -\ln\tanh(2\beta h_0 - \beta h). \tag{22}$$

This family of curves is illustrated by the dotted lines in Fig. 2(b). Each curve is associated with a parameter h_0, which is the value of the field h at which it crosses the self-dual line. The dual of any point on the curve $J_{h_0}(h)$ is easily verified to be $(\tilde{J},\tilde{h}) = (J_{h_0}(2h_0 - h),2h_0 - h)$, which also satisfies (22), consistent with these two points having the same value of $\mathcal{J}$ [recall (21)]. It also follows from (15) that $F_{\rm symm}$ is a symmetric function of $h - h_0$, when evaluated on these lines, just as $F_I(\mathcal{H},\mathcal{J})$ is a symmetric function of $\mathcal{H}$, when evaluated at fixed $\mathcal{J}$.

These curves are useful because the function $F_I(\mathcal{H},0)$ is the cumulant generating function for the magnetization of the Ising model at criticality. One obtains cumulants by taking derivatives with respect to $\mathcal{H}$. Differentiating (20) in this way, evaluation of the right hand side requires $(\partial h/\partial\mathcal{H})_{\mathcal{J}}$ and $(\partial J/\partial\mathcal{H})_{\mathcal{J}}$. In fact, it is sufficient to have the ratio of these derivatives, $(\partial J/\partial h)_{\mathcal{J}}$, but since (22) defines lines of constant

$\mathcal{J}$, this is simply $J'_{h_0}(h) = -2/\sinh(4\beta h_0 - 2\beta h)$. Evaluating this derivative on the self-dual line yields $J'_{h_0}(h_0) = 2\sinh\beta J_0$, where $J_0 = -\ln\tanh h_0$ is the value of the coupling at that point (recall $\sinh 2\beta h_0 \sinh\beta J_0 = 1$ for points on the self-dual line). Hence, we have from (20)

$$\left(\frac{\partial}{\partial\mathcal{H}}\right)^m F_I(0,0) \propto \left(\frac{\partial}{\partial h} + 2\sinh\beta J\frac{\partial}{\partial J}\right)^m F_{\rm symm}(h_c, J_c). \tag{23}$$

The left hand side is proportional to the mth cumulant of the Ising magnetization. All cumulants with odd m are zero by symmetry. An explicit calculation using the symmetry (19) shows that the right hand side is also zero for odd m, as required. [Note also that one expects corrections to (23) due to nonuniversal contributions, as in (20)].

To obtain a more direct interpretation of the right hand side of (23), recall that $(-T\partial_h)^m\beta F_1$ is the mth cumulant of the magnetization $M = \sum_i \sigma_i$ in the TPM, and $(T\partial_J)^m[\beta F_1 + (N\beta J/2)]$ is the mth cumulant of the number of defects, $N_d = \frac{1}{2}\sum_\mu(1-\tau_\mu)$. We therefore define an order parameter for the TPM,

$$\begin{aligned}\mathcal{M} &= -\left(\frac{\partial}{\partial h} + 2\sinh\beta J\frac{\partial}{\partial J}\right)[F_1 + (NJ/2)] \\ &= M - 2N_d\sinh J.\end{aligned} \tag{24}$$

Using (18) and (23), and working at the critical point of the TPM, one sees that cumulants of $\mathcal{M}$ may be written as a sum of terms, one of which is proportional to an Ising model cumulant $(\partial_{\mathcal{H}})^m F_I(0,0)$, while others are nonuniversal corrections. Close to the critical point, the even cumulants of $\mathcal{M}$ show singular behavior while the corrections remain regular. Hence, for systems large enough that the singular terms dominate the regular ones, we expect the critical distribution of $\mathcal{M}$ in the TPM at criticality to match the critical distribution of the Ising magnetization. This prediction will be verified numerically in the following section.

III. NUMERICAL RESULTS FOR THE TPM IN A FIELD

We performed numerical simulations of the TPM in a field to analyze its phase behavior. Working always on the self-dual line (17), we use continuous time Monte Carlo simulations [27,28] to sample a reweighted Boltzmann distribution $P(\sigma) \propto b[M(\sigma)]e^{-\beta E_J(\sigma)+\beta h M(\sigma)}$, where $b(M)$ is a bias function and $M(\sigma) = \sum_i \sigma_i$ is the magnetization. We measure the resulting distribution $P_{\rm b}(M)$ of the magnetization, but we choose the function $b(M)$ so that this sampled distribution does not include any deep minima (free energy barriers) [29]. The "true" distribution $P(M)$ associated with the unbiased model is then easily obtained as $P(M) \propto P_{\rm b}(M)/b(M)$.

The bias potential $b(M)$ is chosen such that $b(M) \approx 1/P(M)$ within a range of M close to its mean. Outside this range we take $b(M)$ independent of M [and $b(M)$ is continuous at the edge of the range]. This means that the sampled distribution $P_{\rm b}(M)$ is approximately constant over a range close to its mean. In practical terms, we start at a relatively high temperature T for which sampling is easy and we collect N_s representative configurations σ^r with $r = 1,2,\ldots,N_s$. We typically take N_s in the range 10^4–10^5: These samples are not fully independent from each other but the sampling runs are long enough that the configurations fully decorrelate within each run. For each sample, we store both the magnetization $M_r = M(\sigma^r)$ and the number of defects $N_{d,r} = N_d(\sigma^r)$ as prescribed by (2). This provides an estimate for a suitable bias potential for further simulations at this temperature,

$$b(M) \propto \frac{1}{\sum_r \delta_{M,M_r}}, \tag{25}$$

where $\delta_{M,M'}$ is the Kronecker δ, so $N_s/b(M)$ is the *empirical distribution* of M. Clearly, (25) can be used only in the range of M for which one has good sampling: For large systems it is also useful to smooth $b(M)$ by forming a histogram with a suitably chosen bin width. We now reduce the temperature to $T - \Delta T$, which corresponds to an increase in J/T of $\Delta_J = J/T - J/(T-\Delta T)$. We also reduce the parameter ε/T by Δ_ε so that the system remains on the self-dual line (16). It is easily verified that, given sufficient sampling,

$$\frac{1}{b_\Delta(M)} \propto \frac{1}{b(M)}\sum_r \delta_{M,M_r} e^{-\Delta_J N_{d,r} - \Delta_\varepsilon M_r} \tag{26}$$

converges to the (unbiased) distribution $P(M)$ at the new state point. We choose a value of ΔT that is small enough for this estimate to be reasonably accurate [essentially, this requires that the exponential weights in (26) do not result in concentration of the probability onto too small a fraction of the samples]. Then $b_\Delta(M)$ is used as a bias potential for a simulation at this new state point. Repeating this procedure allows the temperature to be further reduced. The advantage of the method is that the histogram of M being sampled is approximately flat at each stage. In contrast to unbiased simulations [$b(M) = 1$ for all M], this method is useful when $P(M)$ has two peaks separated by a deep minimum, in which case unbiased calculations tend to sample only from one peak or the other and rarely make transitions between them. However, the flat histogram method facilitates these transitions: See [29] for a review of these kinds of method. We note in passing that this scheme could also be implemented using $\mathcal{M}$ as the order parameter: We would expect similar performance in this case.

Figure 3(a) shows numerical evidence that for large J (and small h), the self-dual line is associated with a first-order phase transition: See also Ref. [17]. The figure shows the distribution $P(m)$ of the magnetization density $m = M/N$ for three state points on the self-dual line. At large J the distribution is bimodal, characteristic of first-order coexistence. A typical configuration at these conditions is shown in Fig. 2(b), showing coexistence of low- and high-m regions, separated by sharp interfaces, as expected for a first-order transition. Based on smaller systems, a previous study [17] speculated that phase separation would not occur for the TPM in a field, but our results show that this does indeed occur if large enough systems are considered. Given the mapping (7), Fig. 3(b) is also a representative configuration of the overlap between two coupled replicas for suitable (J,ε).

As J is decreased (or, equivalently, temperature and field are increased) along the self-dual line, the bimodality in $P(M)$ becomes less pronounced and eventually disappears.

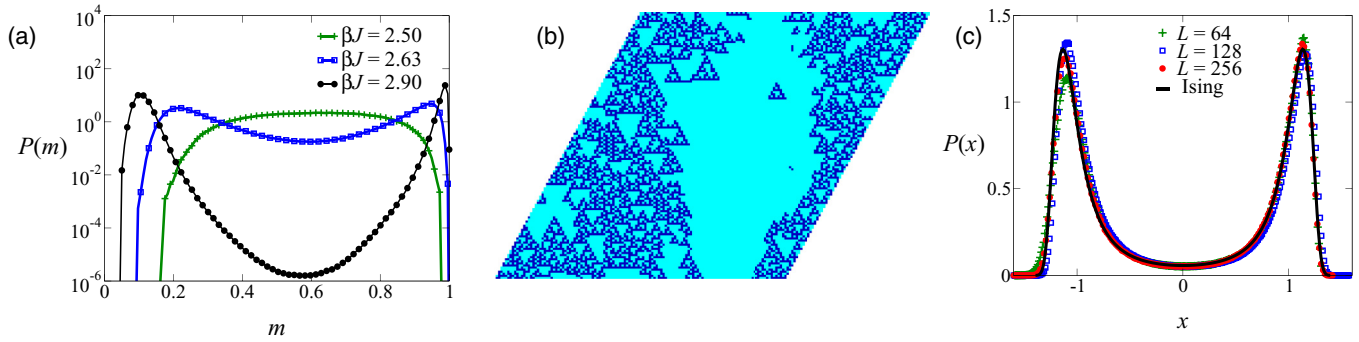

FIG. 3. (Color online) Simulations of the TPM in a field. (a) Distribution of the magnetization at various values of J for state points on the self-dual line (17), at system size $L = 128$. The bimodal distribution $P(m)$ indicates a first-order transition, which disappears on reducing J. From (7), the same distributions would be obtained when considering the overlap between two coupled TPMs, at appropriate state points. (b) Representative configuration at phase coexistence ($\beta J = 2.9$ and $L = 128$) showing interfaces between regions of small and large magnetization (corresponding to regions of small and large overlap in the two-replica problem). (c) At our estimated critical point, ($J_c = 2.634$, $h_c = 0.072$) and for various system sizes, we show distributions of the variable x that is obtained by rescaling the order parameter $\mathcal{M}$ to zero mean and unit variance. The solid line is the corresponding result for the 2D Ising model at criticality [26], indicating that the critical point of the TPM in a field (and therefore of the two coupled TPMs) is in the 2D Ising universality class.

Figure 3(a) indicates that the first-order line terminates at a critical point (J_c,h_c), with $J_c \approx 2.6$. To identify the universality class of this phase transition, we performed a finite-size scaling analysis, using the order parameter $\mathcal{M}$ defined in (24). Note that if all spins are up, one has $M = N$ and $N_d = 0$, giving $\mathcal{M} = N$. On the other hand, in a state with $h = 0$, then $M = 0$ and $N_d = 1/(1 + e^J)$; at low temperatures this gives $\mathcal{M} \approx -N$. In general, one expects a crossover between these two limits as h is increased from 0, with the crossover occurring near $\mathcal{M} = 0$.

As discussed in Sec. II E, one expects the distribution of $\mathcal{M}$ at the critical point to be close to the distribution of the magnetization in a critical Ising model. This provides a sensitive method for identifying the critical point. In order to match the *shape* of the distribution, it is convenient to subtract the mean of $\mathcal{M}$ and rescale so that the distribution has unit variance. We accomplish this by defining $x = (\mathcal{M} - \langle\mathcal{M}\rangle)/\sqrt{\langle(\mathcal{M} - \langle\mathcal{M}\rangle)^2\rangle}$. Working at the critical point, we then expect $P(x)$ to be independent of the system size and that $P(x)$ for the TPM should also match with a similarly rescaled distribution of the magnetization of the 2D Ising model. This allows the critical point to be estimated as the temperature for which $P(x)$ is independent of L and matches with its known Ising form. We find these conditions to be satisfied for a coupling $J/T = 2.634$ (with h/T chosen to be on the dual line). The resulting scaling collapse is shown in Fig. 3(c), which includes data for the TPM at three system sizes, and data for a 2D Ising model at criticality [26]. The data collapse is not perfect but correspondence with the universal (Ising) form is increasingly good as the system size increases, consistent with the singular terms in the free energy becoming dominant as the system size increases (recall Sec. II E). These data confirm the Ising universality class of this phase transition. Note that this collapse implies that cumulant ratios such as $\langle\delta\mathcal{M}^4\rangle_c/\langle\delta\mathcal{M}^2\rangle^2$ must take the appropriate universal values, independent of system size. Hence, the data in Fig. 3(c) is sufficient to ensure that cumulant ratio crossings (see, for example, [16]) also take place at this estimated critical temperature. (Note also that while the critical distribution of the order parameter is bimodal, the transition is *second order* [26]: The separation of the peaks vanishes as the system size $L \to \infty$ and the depth of the trough between them remains constant, in contrast to first-order transitions for which the peak spacing remains constant and $P(M) \to 0$ in the trough.)

We also calculated the ratio of the susceptibilities $\chi = L^{-d}\langle\delta\mathcal{M}^2\rangle$ at criticality, for the system sizes $L = 128$ and $L = 256$. We find $\chi(L = 256)/\chi(L = 128) = 3.36$. Theory predicts that this ratio should scale as $L^{\gamma/\nu}$, where $(\gamma,\nu) = (7/4,1)$ are the susceptibility and correlation length exponents for 2D Ising universality. This yields a prediction of $2^{7/4} \approx 3.364$ for the ratio of susceptibilities, consistent with our results.

Bringing together these results, we arrive at the phase diagram shown in Fig. 4. We reintroduce the temperature $T = \beta^{-1}$ as an explicit parameter and plot the phase diagrams as a function of T/J and ε/J since this is conventional representation in supercooled (glassy) liquids. The form of this phase diagram was proposed in [18]. However, the results here now provide both the correct location and universality class of the critical point as compared to that work, whose arguments were based on the incorrect assumption that the TPM in a field was equivalent to the generalized Baxter Wu model [30]. Our finding that the critical point is in the Ising class is interesting, since this is the expected result from other general arguments [23] and what is observed in simulation of coupled liquids [24,25].

A key feature of the plaquette models is that the first-order transition line does not intersect the $\varepsilon = 0$ axis except at $T = 0$ [18]. However, the duality line at low temperatures is $\beta\varepsilon \simeq 2e^{-\beta J}$, so the field $\varepsilon = \varepsilon^*$ at which the first-order phase transition takes place is very small. Indeed, simple extrapolation of the first-order line from high-temperature data might lead one to propose a phase transition at $\varepsilon = 0$ and some $T = T_K > 0$, as predicted in mean-field theories [6]. In this case the scenario shown in Fig. 4 might be hard to distinguish from the mean-field picture. However, direct simulations near this proposed phase transition (if possible) would demonstrate that there are no diverging fluctuations near the first-order

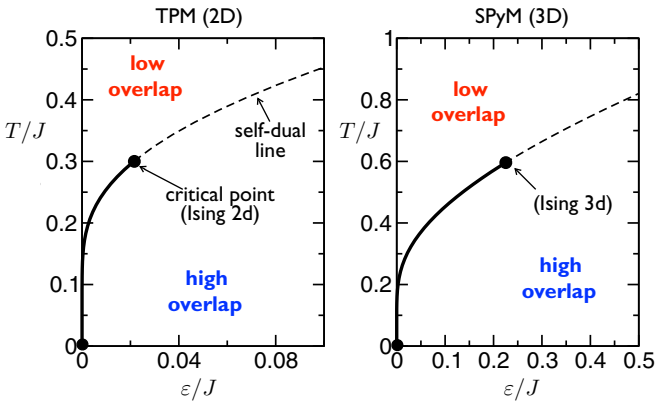

FIG. 4. (Color online) Phase diagrams of coupled plaquette models, the 2D TPM (left) and the 3D SPyM (right, see also Sec. V below). The solid line corresponds to a line of first-order transitions between a thermodynamic phase of small overlap and one of large overlap between the replicas. This curve is on the self-dual line (17) (dashed line). The first-order transition line ends at a critical point that is in the 2D Ising universality class for the TPM and the 3D Ising universality class for the SPyM. (For the 3D model in a field, the location of the critical point and the form of the phase diagram were already calculated in [16], so we do not repeat this calculation here. The results shown here for the coupled replicas then follow from the mapping in Sec. II C.)

transition line, in contrast to the mean-field scenario where length and time scales diverge at T_K.

IV. ACTIVE-INACTIVE DYNAMICAL TRANSITIONS IN THE TPM

The TPM falls into a category of glassy models that are thermodynamically simple but where glassy behavior arises because of nontrivial dynamical pathways to the equilibrium state at low temperature [11,12,15,31]. In fact, the dynamics of the TPM [15] is closely related to that of a two-dimensional East model [11–13]. Many kinetically constrained models, including the East model, display dynamical phase transitions—phase transitions in the space of trajectories—between a phase with a high dynamical activity, K, and one with low dynamical activity [8].

These lattice models evolve in time by flips of spin variables (for plaquette models, these are the σ_i variables). Each time a spin changes its state, this changes the configuration of the system: The activity K is defined as the total number of configuration changes in a trajectory [32,33]. By coupling a field s to the dynamical activity, one can define the so-called s ensemble (also known as the exponentially biased or tilted ensemble), where the probability of obtaining a trajectory X_τ of total time extension τ is reweighted by its activity [8,9,32],

$$P_s(X_\tau) = \frac{e^{-sK} P_0(X_\tau)}{Z_s(\tau)}. \tag{27}$$

Here $P_s(X_\tau)$ is the probability of the trajectory X_τ in the s ensemble, $P_0(X_\tau)$ is the unbiased probability of this trajectory (the one generated by the actual dynamics of the system), and $Z_s(\tau)$ is the moment generating function of the activity K, in effect a partition sum for trajectories. For long times, $Z_s(\tau)$ takes on a large-deviation form [8,32,34],

$$Z_s(\tau) \propto e^{-\tau\theta(s)}, \tag{28}$$

where $\theta(s)$ plays the role of a dynamical free-energy and is the scaled cumulant generating function of the activity. The mean time-intensive activity of trajectories in the s ensemble,

$$k(s) \equiv \langle K \rangle_s / \tau, \tag{29}$$

where the average is with respect to the ensemble (27), can thus be obtained from $k(s) = \frac{d}{ds}\theta(s)$. Dynamical—or space-time—phase transitions manifest as singularities in $\theta(s)$ [8,9,32,35]. We also define the susceptibility

$$\chi(s) \equiv -\frac{dk}{ds} = \tau^{-1}\langle \delta K^2 \rangle_s. \tag{30}$$

Like the East model, the dynamical relaxation of the TPM is hierarchical, due to energy barriers to relaxation that are logarithmic in the linear size of relaxing regions [12]. This, in turn, leads to a "parabolic" [36] super-Arrhenius law for the typical relaxation time in equilibrium as a function of temperature. Given the similar dynamical properties of the TPM and the East model dynamics, a natural question is whether the TPM also displays active-inactive space-time transitions.

In order to answer this question numerically, we make use of transition path sampling (TPS) [37] to efficiently sample trajectories in the s ensemble. For the purposes of numerical efficiency, we exploit the fact that different trajectory ensembles can be defined by fixing different dynamical quantities, such that for long trajectories these distinct ensembles become equivalent [38,39]. This is analogous to ensemble equivalence in the thermodynamic limit of standard statistical mechanics. We consider, in particular, the x ensemble introduced in [39], that is, the ensemble of trajectories with fixed number of configuration changes, i.e., activity, K, but where the overall trajectory time extension τ fluctuates. In contrast, the s ensemble above is one where trajectories are of fixed overall time but their activity fluctuates. For large K and τ these ensembles can be shown to be equivalent [39,40]. For the case of the TPM the x ensemble is particularly efficient to simulate (see [39] and [41] for details) and the functions $k(s)$ and $\chi(s)$ can be recovered from the ensemble equivalence.

Figure 5 shows the results of the s-ensemble analysis of the TPM. It shows the average activity for a system of size $N = 8 \times 8$ as a function of s, at temperature $T = 0.5$ (note this is the TPM in the absence of field). As the length of the trajectories is increased the change in $\langle K \rangle_s$ becomes more pronounced, as seen in the corresponding susceptibilities. This is indicative of a first-order transition at some $s_c \gtrsim 0$. Similar size scaling is observed by changing system size, as shown in the insets. The dependence of this transition on the temperature is discussed in Sec. V C, together with similar results for the (3D) SPyM.

We note at this point that these transitions are inherently dynamical in nature. For example, one might also consider a plaquette model that evolves in time by dynamical rules in which single plaquette variables may change their state independently. In this case the dynamics is that of a (grand canonical) free gas of defective plaquettes and the statistics of K are simply those of a Poisson process (the same situation

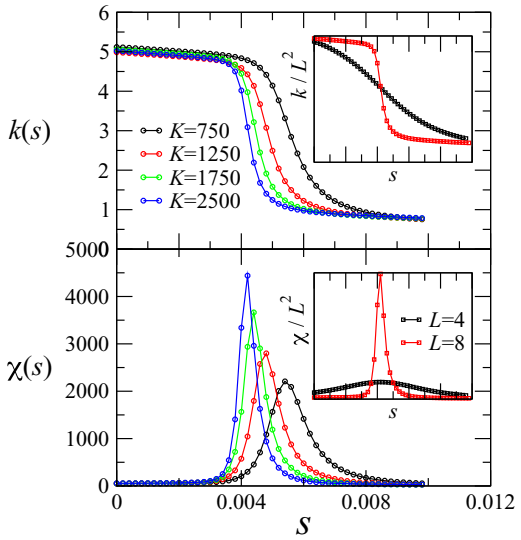

FIG. 5. (Color online) Average activity $k(s)$ and the associated susceptibility $\chi(s)$ in the TPM at $T = 0.5$. The main panels show data for system size $L = 8$. These results were obtained by the x-ensemble method (see text), using trajectories with fixed numbers of events K, as shown. On increasing the trajectory length, the crossover from active to inactive behavior becomes increasingly sharp and the susceptibility peak increases. The behavior for smaller systems ($L = 4$) is shown in the insets, with both quantities normalized by the system size. In the absence of a phase transition, one expects both $k(s)/L^d$ and $\chi(s)/L^d$ to be independent of L, so the sharper crossover at $L = 8$ is again consistent with an underlying phase transition.

is observed when considering kinetically constrained models from which the constraint has been removed [8]). Similarly, the transition in the coupled-replica system of the previous section is also destroyed if one measures the overlap at the defect level [that is, one replaces Q in (4) by $Q_d = \sum_\mu \tau_\mu \tau'_\mu$]. In summary, observation of the phase transitions discussed in this paper requires a suitable choice of order parameter (Q not Q_d), and the dynamical phase transitions also depend on the dynamical rules by which the systems evolve.

V. OVERLAP AND ACTIVITY TRANSITIONS IN A THREE-DIMENSIONAL PLAQUETTE MODEL

In order to explore whether the static and dynamical transitions found above for the TPM are present in dimensions other than two, it is of interest to generalize the TPM to higher dimensions. One of the reasons is that if one wishes to use plaquette models to study "quenched" coupled replicas [6,23,24] or "random pinning" [42–44], the distinction between two and three dimensions may be very significant, due to the inability of two-dimensional systems in random fields to support first-order static transitions [45]. Here we introduce a three-dimensional model that is similar to the TPM.

A. Model

The model we consider is defined on a 3D bcc lattice. The "plaquettes" are upward-pointing square-based pyramids, each containing five spins. Considering the standard bcc unit cell, one such pyramid is formed by the spin at the center

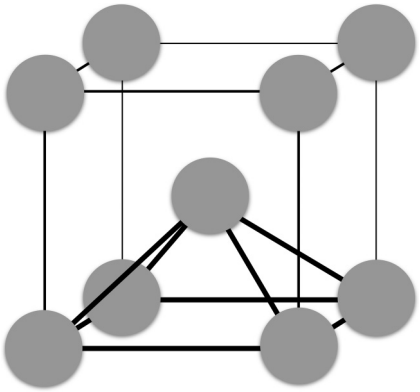

FIG. 6. The SPyM consists of spins (gray circles) on the sites of a bcc lattice which interact in quintuplets at the vertices of upward-pointing square pyramids. One such pyramid is indicated; The central spin also participates in four other upward-pointing pyramids whose apexes are the four spins on the upper face of the cube.

of the cube together with the four spins at the corners of the lower face; see Fig. 6. We call this model the square pyramid plaquette model, or SPyM. Its energy function reads

$$E_J(\sigma) \equiv -\frac{J}{2} \sum_\mu \sigma_{i_\mu} \sigma_{j_\mu} \sigma_{k_\mu} \sigma_{l_\mu} \sigma_{m_\mu} \quad \text{(SPyM)}, \tag{31}$$

where μ runs over all the pyramidal plaquettes on the lattice, and the location of the five interacting spins $\sigma_{i_\mu} \cdots \sigma_{m_\mu}$ is shown in Fig. 6. This is "model 1" of [16].

Just like the TPM, the SPyM has a one-to-one correspondence between spin and plaquette configurations. An alternative model [10,16] may be defined on a face-centered cubic lattice, in which the plaquettes are tetrahedral pyramids ("model 2" of [16]). However, in this case each interaction involves four spins, so the system has a global spin-flip symmetry, and the spin-plaquette correspondence is not exact in finite (periodic) systems. However, these deviations from the one-to-one correspondence are irrelevant in the thermodynamic limit.

Returning to the SPyM, we explicitly demonstrate the one-to-one correspondence between spins and plaquettes by a general method that applies also to the TPM. The total number of configurations of the spin variables is 2^N, and this is also equal to the total number of configurations of the plaquette variables, by construction of the model. For any spin configuration, the configuration of the plaquette variables is uniquely specified since the τ_μ are defined in terms of the spins σ_i. However, it might be that some plaquette configurations can be achieved by more than one spin configuration, in which case the mapping is not one to one (this happens for example in the square plaquette model [20] with periodic boundary conditions). In that case the equal numbers of spin and plaquette configurations means that there must exist plaquette configurations that cannot be realized by any realization of the spin variables. To rule this out and establish the one-to-one mapping, we now show how a spin configuration may be constructed for any given plaquette configuration.

We choose as basis vectors for the lattice $\vec{a}_1 = (1,0,0)$, $\vec{a}_2 = (0,1,0)$, and $\vec{a}_3 = (-1,-1,\sqrt{2})/2$. We focus on systems

whose sites are at $l\vec{a}_1 + m\vec{a}_2 - n\vec{a}_3$ with $l,m,n \in \{0,1,2,\ldots,L-1\}$, with periodic boundaries (so, for example, sites with $n = L-1$ are neighbors of those with $n = 0$). We indicate the location of the μth pyramid by the position of the spin at the apex. The plaquette variable τ_μ for $\mu = (i,j,k)$ is then

$$\tau_{(i,j,k)} = \sigma_{(i,j,k)}\sigma_{(i,j,k-1)}\sigma_{(i-1,j,k-1)}\sigma_{(i,j-1,k-1)}\sigma_{(i-1,j-1,k-1)}. \tag{32}$$

Following the same reasoning as in [10], we can invert this relation in terms of a "Pascal pyramid": The idea is to demonstrate that introducing a single defect into the system corresponds to flipping a particular set of spins.

Starting from the ground state, we demonstrate the procedure by introducing a single defect at the origin: This affects those spins in upper layers which lie on the sites of an inverted Pascal pyramid (or fractal pyramid). Assuming that the central spin in Fig. 6 is at the origin, we flip that spin, which introduces a defect in the pyramid below it. In order to avoid any other defects, we also flip the four spins on the top face of the cube shown in Fig. 6, which ensures that there are no defects in any of the pyramids pointing upward from the origin. Iterating this procedure for all other layers, the final spin configuration is

$$\sigma_{(i,j,k)} = 1 - 2\left[\binom{k}{i}\binom{k}{j} \mod 2\right], \tag{33}$$

where $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ are combinatorial numbers, and $0 \leqslant i,j \leqslant k$ (all other spins $\sigma_{(i,j,k)} = 1$). Given periodic boundary conditions, this procedure determines all spins in the system: On setting the final layer of spins, it may be that defects in the final layer are unavoidable. However, for systems whose linear size L is a power of 2, it is easily shown that this procedure produces a final state with exactly one defect.

Now observe that for any spin configuration, flipping the set of spins for which $\sigma_{(i,j,k)} = -1$ in (33) inverts the state of the plaquette variable just below the origin, leaving all other plaquette variables constant. Similarly, to flip the state of any other plaquette, one applies a spatial translation to the same set of spins and flips all the spins within this translated set. Hence, by repeatedly applying this procedure, one can generate a spin configuration that corresponds to any given configuration of the plaquette variables. This establishes the one-to-one correspondence between spin and plaquette configurations by the argument given above.

This correspondence between spin and defect configurations means that the thermodynamics of the SPyM is that of a free binary gas of plaquettes. Furthermore, the relaxational dynamics is similar to that of a (3D) East model. Figure 7 shows the decay of the energy at low temperatures starting from a $T = \infty$ configuration. We see the characteristic hierarchical decay of both the East model and the TPM: The energy decays in steps with characteristic time scales $\tau_n = e^{n\beta J}$, with $n = 0,1,2,\ldots$. These steps become apparent when plotting as a function of the rescaled time variable $(T/J)\ln t$ [46]. The inset to Fig. 7 shows that the equilibrium relaxation time of the SPyM is super-Arrhenius, as in the East model and the TPM.

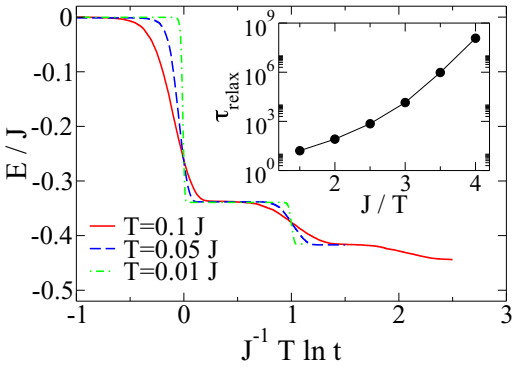

FIG. 7. (Color online) Relaxation of the energy of the SPyM at low temperature starting from a random configuration (the system size is $L = 16$). The curve shows the characteristic plateaus indicative of hierarchical relaxation, as in the East model and the TPM. (Inset) Average relaxation time as a function of inverse temperature, showing super-Arrhenius behavior.

B. Phase transition in (annealed) coupled replicas

The SPyM possesses the exact duality described in Sec. II C. In particular, the properties of the SPyM in a field were studied in Ref. [16] ("model 1" of that paper), where it was found that on the self-dual line there is a first-order transition between phases of small and large magnetization terminating at a critical point in the 3D Ising class. From those results we can directly infer the phase diagram of the two coupled SPyMs via the mapping of Sec. II C. The result is shown in Fig. 4. This phase diagram is similar to that of the TPM, except that the range of phase coexistence is larger and the critical point occurs at higher temperature.

C. Evidence for a dynamical (space-time) phase transition

As well as the phase transition for coupled replicas in the SPyM, we also present evidence for a space-time phase transition, similar to that shown for the TPM in Fig. 5. The results for the SPyM are shown in Fig. 8, for temperature $T/J = 0.65$ and linear system size $L = 4$. There is good evidence for a sharp transition at $s = s^* > 0$, as found in the TPM.

In Figs. 8(b) and 8(c), we show how the crossover in activity varies with the temperature T/J, for both the TPM and the SPyM. Simple estimates [47,48] indicate that if the inactive state is metastable and relaxes to equilibrium via some kind of nucleation process with rate $\gamma_{\rm nuc}$ per unit volume, then $s^* \approx \gamma_{\rm nuc}/\delta k$, where δk is the activity difference (per unit space-time) between the active and inactive states [47]. We attribute the existence of the transition in this model to a stable inactive state with almost no defects. We expect that the rate for relaxation back to equilibrium is a strongly decreasing function of temperature, which is consistent with the increasing s^* as T/J increases. (The activity difference Δk between active and inactive states increases with T/J, but this dependence is much weaker than that of the relaxation rate.)

Comparing Figs. 4 and 8, natural questions concern whether the crossover lines identified in Figs. 8(b) and 8(c) are indeed first-order phase transition lines and whether the true phase behavior involves a first-order line for low temperature that

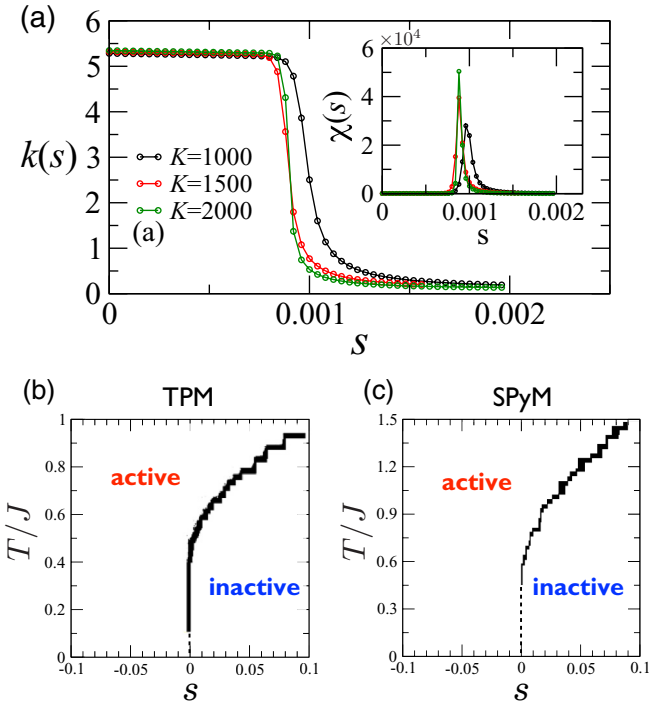

FIG. 8. (Color online) (a) Activity as a function of s in the SPyM for system size $L = 4$, various trajectory lengths, and $T/J = 0.65$. The inset shows the corresponding susceptibility. (b), (c) s-ensemble phase diagrams for the TPM and SPyM. The solid curves are an estimate of the transition point from the simulations. The dashed lines are extrapolations in the low-temperature regime inaccessible to numerics.

terminates at a critical point. Our numerical results are not sufficient to answer this question, due to the significant computational expense of sampling the s ensemble. However, we expect on general grounds that the situation should be similar to the softened Fredrickson-Andersen model considered in [19], in which case the first-order line would indeed end at a finite-temperature critical point. This scenario is consistent with the results presented here.

VI. DISCUSSION

A. Connection of phase transitions to long-lived metastable states

The phase behavior shown in Figs. 4 and 8 reveals striking similarities between thermodynamic transitions (for coupled replicas) and dynamic transitions (based on dynamical activity). The two transitions are distinct and we do not believe that they are related by any exact mapping. Nevertheless, we argue in the following that these transitions are connected to the existence of long-lived metastable states, which are intrinsically linked to the glassy behavior in these systems.

Consider first transitions for annealed replicas. If we work at the phase coexistence point, but within the high-overlap phase, the system occupies low-energy states. For $\varepsilon = 0$, these states would be metastable (with finite lifetimes). This metastability means that when localized low-overlap regions are generated by thermal fluctuations, these regions tend to shrink, just as small fluctuations tend to shrink within classical nucleation theory. Escape from the metastable state requires a collective process that operates on some finite length scale. As ε is increased from zero, this length scale increases, as does the associated free energy barrier: Both diverge at the coexistence point where the high-overlap phase becomes stable.

The situation for dynamical phase transitions is similar, except that one should think of trajectories of the system as $(d+1)$-dimensional objects that exist in space-time. If one works at the dynamical phase coexistence point (some $s = s^* > 0$), then inactive trajectories dominate the s ensemble. During these trajectories, the system remains localized in low-energy metastable states, with small thermal fluctuations of the activity, associated with space-time "bubbles" [7]. If the trajectory length τ is less than the time required for escape from the metastable state, similar inactive trajectories can be generated with unbiased ($s = 0$) dynamics by taking initial conditions from low-energy metastable states. Transformation of such a trajectory into a *typical* equilibrium trajectory involves the introduction of an active space-time bubble, which subsequently grows to macroscopic size. The connection with metastability arises because if one introduces a small active bubble within an inactive trajectory, one expects to incur a cost in probability (if this were not the case, the state would not be metastable since it would readily relax back to equilibrium). As in the case of overlap fluctuations, the critical bubble size and the probability barrier increase as s is increased from zero, diverging at the coexistence point.

We argue that this analogy between phase coexistence phenomena induced by s and ε fields provides a qualitative explanation of the similarity between Figs. 4 and 8, in that both are linked to the existence of metastable states that can be observed in unbiased ($s = 0 = \varepsilon$) systems. The relevant metastable states have low energy: Both the inactive state of Fig. 8 and the high-overlap state of Fig. 4 have much lower defect concentrations than the equilibrium average value $c \approx e^{-\beta J}$. For large s, the system minimizes its propensity for dynamical activity by removing defects, so that $N_d/N \approx 0$; for large coupling ε it is easy to show that $N_d/N \approx e^{-2\beta J}$ since the system has an effective temperature $1/(2\beta)$ [49]. [Considering the partition function (5), for large ϵ we have $\sigma^a \approx \sigma^b$, in which case the Boltzmann factor in (5) reduces to $e^{-2\beta E_J(\sigma^a)+\varepsilon N}$, and the distribution over configurations σ^a is equilibriumlike at temperature $1/(2\beta)$. The general case considered in [49] involves m coupled replicas in the case where the coupling is strong enough to localize all replicas in the same metastable state, but weak enough that it has a negligible effect on intrastate fluctuations.]

Evidence for phase coexistence induced by s and ε fields have both been presented in atomistic models [9,24]. By contrast, in kinetically constrained models (KCMs), dynamical phase coexistence can be induced by the s field [8] but there is no such transition as a function of ε. Metastable states can be identified in KCMs [48]; here a *state* is defined [50] as a region of configuration space for which the equilibration time within the state is much smaller than the time to escape from it. (As argued in [50], this is a robust definition in finite-dimensional systems at $\varepsilon = 0$, for which mean-field constructions break down.) Given this definition of metastability, phase coexistence at positive s may be expected,

as argued above. However, this metastability does not lead to phase coexistence for any ε: In fact, the free energy for two coupled KCMs can be obtained exactly and the statistics of the overlap are trivial in this case. Clearly, the static construction based on coupled replicas does not reveal the metastability in this model: The reason is that the kinetic constraints generate metastable behavior by reducing the ability of the system to evolve from high-overlap to low-overlap states. (To see this, note that if one removes the kinetic constraints, the metastable states disappear but the results for coupled replicas do not change.)

A similar situation occurs when considering plaquette models with different overlap measures and different rules for their dynamical evolution. As discussed in Sec. IV, changing the dynamical rules of the plaquette models so that plaquette (defect) variables flip independently destroys the dynamical phase transition. (Viewing the plaquette models in a defect representation, this is equivalent to removing a kinetic constraint [12].) Also, changing the definition of overlap from Q to $Q_d = \sum_\mu \tau_\mu \tau'_\mu$ means that the phase transition for the coupled replicas is lost. (This also happens in plaquette models on random graphs for which exact mean-field calculations are possible [22].) The general point here is that the natural dynamical definition of metastability [50] coincides with the presence of phase transitions for coupled replicas only if the natural dynamics of the system leads to unconstrained local changes in the overlap. So if the dynamics has (unconstrained) single spin flips, then the natural overlap parameter is the spin overlap and one observes both static and dynamic transitions in plaquette models. If the dynamics has independent changes to defect variables and one uses the defect overlap, then both static and dynamical transitions are lost, so the two constructions are still consistent with each other. If one uses spin-flip dynamics and an overlap based on defects then the dynamical transition survives but the coupled replicas are insensitive to it. (This is similar to the KCM case: The dynamical rules for defect variables in plaquette models are constrained, but the spin-overlap is not sensitive to this constraint.) Finally, if one uses independent-defect dynamics but an overlap based on spins, then the dynamical transition is lost but the static one survives: The dynamical and static constructions probe different aspects of the system in this case and give different results.

We end this section by noting that the inactive and high-overlap states in Figs. 4 and 8 are structurally distinct from equilibrium states at temperature T, and also from the active and low-overlap states. For example, as noted above, configurations deep inside the high-overlap phase have energies and spin correlation functions representative of equilibrium at temperature $T/2$, while the low-overlap phase is close to an equilibrium state at temperature T. The states observed at finite ε are not exactly the equilibrium states at temperature T or $T/2$ that are found for $\varepsilon = 0$ or $\varepsilon \to \infty$, but their properties are qualitatively similar, and there is a significant jump in the energy $E_J(\sigma)$ on crossing the first-order transition line. This is quite different from ensembles with a quenched coupling between replicas [51,52], where the structures of high- and low-overlap states should be statistically (almost) indistinguishable, with (at most) a small discontinuity in $E_J(\sigma)$ at the transition (see also [42]).

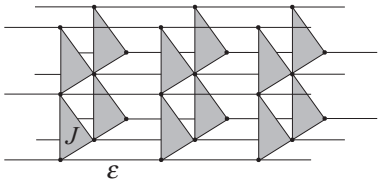

FIG. 9. Three-dimensional stack of coupled two-dimensional TPMs.

B. Connection between multiple coupled replicas and biased-activity ensembles

The connection between Figs. 4 and 8 can be further motivated through a generalization of the coupled two-replica system discussed above. Given a plaquette model in dimension d, consider the associated $d+1$ system composed of many replicas of the d-dimensional system arranged parallel to each other along the extra dimension; see Fig. 9. Such system of n coupled replicas has an energy

$$E_n(\sigma^1,\sigma^2,\dots,) \equiv E_J(\sigma^1) + E_J(\sigma^2) + \cdots - \varepsilon \sum_i \left(\sigma_i^1\sigma_i^2 + \sigma_i^2\sigma_i^3 + \cdots\right). \quad (34)$$

Using the methods of Refs. [17,18,30] it is easy to prove that the partition sum of the $(d+1)$-dimensional problem also has an exact duality,

$$Z_n(J,\varepsilon) = (\sinh\beta J \sinh\beta\varepsilon)^{Nn} Z_n(J^*,\varepsilon^*), \quad (35)$$

where we have assumed periodic boundary conditions in the transverse direction, and (J,ε) and (J^*,ε^*) are related again by (14). Similar results were found in [53] for other classes of plaquette models. Given this duality, we expect the phenomenology of this many-replica system to be similar to that of two replicas, except that any phase transitions should be in the $(d+1)$-dimensional Ising universality class (assuming that both longitudinal and transverse dimensions are taken to infinity in the thermodynamic limit).

The partition sum Z_n has a natural transfer matrix representation in the transverse direction, $Z_n = \mathrm{Tr}(\mathbb{T}^n)$, with

$$\mathbb{T} = \cosh^N(\beta J/2) e^{N\beta\epsilon} \bigotimes_\mu \left[1 + \tanh(\beta J/2)\sigma^z_{i_\mu}\sigma^z_{j_\mu}\cdots\sigma^z_{k_\mu}\right] \bigotimes_i \left(1 + e^{-2\beta\epsilon}\sigma_i^x\right), \quad (36)$$

where $\sigma^{x,z}$ are Pauli matrices. This, in turn, can be related to the generator of (imaginary time) quantum evolution in the usual manner [53] when the transverse coupling is large and the longitudinal one small [$e^{-2\beta\varepsilon}, \tanh(\beta J/2) \ll 1$], so that $Z_n \propto \exp(-tH)$, with

$$H \equiv -h\sum_i \sigma_i^x - g\sum_\mu \sigma^z_{i_\mu}\sigma^z_{j_\mu}\cdots\sigma^z_{k_\mu}, \quad (37)$$

where

$$\delta t h = e^{-2\beta\varepsilon}, \quad \delta t g = \beta J/2, \quad t = n\delta t. \quad (38)$$

The Hamiltonian (37) generates dynamics in the transverse direction. While it is not derived from a stochastic operator, it has the basic features of the generator [8,32] for the

dynamical ensemble defined in (27): an off-diagonal part (the σ^x terms) that performs configuration changes and a diagonal part (the σ^z) plaquette terms associated with the escape rate. The parameter s in the s-ensemble operator controls the relative strength of the diagonal and off-diagonal terms [8,32], in analogy with the balance between h and g in (37). Furthermore, the duality (35) implies a duality $h \leftrightarrow g$ in (37), with the possibility of a dynamical transition at that self-dual point $g = h$. This connection between a static transition in the $d+1$-dimensional problem (34) (itself closely connected to the static transition in the two-replica plaquette system) and a dynamical transition in the d-dimensional system (37) provide another rationalization of the similarities between Figs. 4 and 8.

C. Outlook

Plaquette spin models have several features that make them attractive for studies of the glass transition. As we have shown here, exact results can be derived, which guide numerical studies of phase behavior and many-body correlations. The models are also computationally much less demanding than atomistic models of supercooled liquids, so that (for example) finite-size scaling over a large range of system sizes can be performed to analyze phase transitions. The equilibrium relaxation of the models follows a dynamical facilitation scenario, in which point defects play a central role. However, there are strong many-body correlations, and the statistics of overlap fluctuations are rich and complex, as anticipated in the theory of Franz and Parisi [6]. In this sense, the models provide a bridge between different theories. Indeed, as argued in [20], one might describe plaquette models by a modified form of RFOT, but with two important caveats: (i) the analog of the Kauzmann transition occurs at zero temperature in these models and (ii) the interfacial cost associated with growing droplets of a new state within a typical equilibrium state scales logarithmically in the droplet size (not as a power law, as anticipated by RFOT).

Looking forward, we hope that further work on plaquette models (particularly in $d = 3$) will show to what extent mean-field [6] and RFOT ideas can be modified to apply in this setting. We can imagine that the apparently different physical pictures envisaged by thermodynamic and dynamical theories of the glass transition [4] might *both* be applicable in these models. In that case, it is not clear whether some new results would be required to discriminate between the theories or whether they might, in fact, offer complementary descriptions of the same phenomena.

ACKNOWLEDGMENTS

We thank Nigel Wilding for helpful discussions and for providing the Ising model data for Fig. 3. This work was supported in part by Engineering and Physical Sciences Research Council Grant No. EP/I017828/1. R.L.J. was supported by the EPRSC through Grant No. EP/I003797/1.

[1] M. D. Ediger, C. A. Angell, and S. R. Nagel, J. Phys. Chem. **100**, 13200 (1996).
[2] A. Cavagna, Phys. Rep. **476**, 51 (2009).
[3] L. Berthier and G. Biroli, Rev. Mod. Phys. **83**, 587 (2011).
[4] G. Biroli and J. P. Garrahan, J. Chem. Phys. **138**, 12A301 (2013).
[5] V. Lubchenko and P. G. Wolynes, Annu. Rev. Phys. Chem. **58**, 235 (2007).
[6] S. Franz and G. Parisi, Phys. Rev. Lett. **79**, 2486 (1997).
[7] D. Chandler and J. P. Garrahan, Annu. Rev. Phys. Chem. **61**, 191 (2010).
[8] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, Phys. Rev. Lett. **98**, 195702 (2007).
[9] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, Science **323**, 1309 (2009).
[10] J. P. Garrahan, J. Phys. Condens. Matter **14**, 1571 (2002).
[11] J. Jäckle and S. Eisinger, Z. Phys. B **84**, 115 (1991).
[12] F. Ritort and P. Sollich, Adv. Phys. **52**, 219 (2003).
[13] J. Garrahan and D. Chandler, Proc. Natl. Acad. Sci. USA **100**, 9710 (2003).
[14] M. E. J. Newman and C. Moore, Phys. Rev. E **60**, 5068 (1999).
[15] J. P. Garrahan and M. E. J. Newman, Phys. Rev. E **62**, 7670 (2000).
[16] J. R. Heringa, H. W. J. Blöte, and A. Hoogland, Phys. Rev. Lett. **63**, 1546 (1989).
[17] S. Sasa, J. Phys. A **43**, 465002 (2010).
[18] J. P. Garrahan, Phys. Rev. E **89**, 030301 (2014).
[19] Y. S. Elmatad, R. L. Jack, D. Chandler, and J. P. Garrahan, Proc. Natl. Acad. Sci. USA **107**, 12793 (2010).
[20] R. L. Jack and J. P. Garrahan, J. Chem. Phys. **123**, 164508 (2005).
[21] If the plaquette interactions do not allow for a spin-defect duality, as for example with spins in a cubic lattice with interactions on the square faces (rather than on the cubes), then the static properties may be nontrivial. See, for example, D. Johnston, J. Phys. A **45**, 405001 (2012), and references therein.
[22] This is even true in mean-field versions of plaquette modelsB-semi see L. Foini, F. Krzakala, and F. Zamponi, J. Stat. Mech. (2012) P06013.
[23] S. Franz and G. Parisi, J. Stat. Mech. (2013) P11012.
[24] L. Berthier, Phys. Rev. E **88**, 022313 (2013).
[25] G. Parisi and B. Seoane, Phys. Rev. E **89**, 022309 (2014).
[26] D. Nicolaides and A. D. Bruce, J. Phys. A **21**, 233 (1988).
[27] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comp. Phys. **17**, 10 (1975).
[28] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, Oxford, U.K., 1999).
[29] A. D. Bruce and N. B. Wilding, Adv. Chem. Phys. **127**, 1 (2003).
[30] Y. Deng, W. Guo, J. R. Heringa, H. W. Blöte, and B. Nienhuis, Nucl. Phys. B **827**, 406 (2010).
[31] G. H. Fredrickson and H. C. Andersen, Phys. Rev. Lett. **53**, 1244 (1984).
[32] V. Lecomte, C. Appert-Rolland, and F. van Wijland, J. Stat. Phys. **127**, 51 (2007).

[33] M. Baiesi, C. Maes, and B. Wynants, Phys. Rev. Lett. **103**, 010602 (2009).
[34] H. Touchette, Phys. Rep. **478**, 1 (2009).
[35] T. Speck, A. Malins, and C. P. Royall, Phys. Rev. Lett. **109**, 195703 (2012).
[36] Y. S. Elmatad, D. Chandler, and J. P. Garrahan, J. Phys. Chem. B **113**, 5563 (2009).
[37] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[38] R. Chetrite and H. Touchette, Phys. Rev. Lett. **111**, 120601 (2013).
[39] A. A. Budini, R. M. Turner, and J. P. Garrahan, J. Stat. Mech. (2014) P03012.
[40] J. Kiukas, M. Guţă, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. E **92**, 012132 (2015).
[41] R. M. Turner, T. Speck, and J. P. Garrahan, J. Stat. Mech.: Theory Exp. (2014) P09017.
[42] C. Cammarota and G. Biroli, Proc. Natl. Acad. Sci. USA **109**, 8850 (2012).
[43] R. L. Jack and L. Berthier, Phys. Rev. E **85**, 021120 (2012).
[44] W. Kob and L. Berthier, Phys. Rev. Lett. **110**, 245702 (2013).
[45] M. Aizenman and J. Wehr, Phys. Rev. Lett. **62**, 2503 (1989).
[46] P. Sollich and M. R. Evans, Phys. Rev. Lett. **83**, 3238 (1999).
[47] R. L. Jack, L. O. Hedges, J. P. Garrahan, and D. Chandler, Phys. Rev. Lett. **107**, 275702 (2011).
[48] R. L. Jack, Phys. Rev. E **88**, 062113 (2013).
[49] R. Monasson, Phys. Rev. Lett. **75**, 2847 (1995).
[50] J. Kurchan and D. Levine, J. Phys. A **44**, 035001 (2011).
[51] C. Cammarota, A. Cavagna, I. Giardina, G. Gradenigo, T. S. Grigera, G. Parisi, and P. Verrocchio, Phys. Rev. Lett. **105**, 055703 (2010).
[52] L. Berthier and R. L. Jack, Phys. Rev. Lett. **114**, 205701 (2015).
[53] C. Xu and J. E. Moore, Phys. Rev. Lett. **93**, 047003 (2004); Nucl. Phys. B **716**, 487 (2005).